DOCUMENT RESUME

ED 342 803 TM 017 971

AUTHOR McMorris, Robert F.; And Others
TITLE Why Do Young Students Change Answers on Tests?
PUB DATE 91
NOTE 22p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Achievement Gains; Behavior Patterns; *Elementary School Students; Grade 5; Grade 6; High Achievement; Intermediate Grades; Interviews; *Scores; Sex Differences; *Student Attitudes; *Test Wiseness
IDENTIFIERS *Answer Changing (Tests)

ABSTRACT

Effects of answer changing on tests by 133 fifth- and sixth-grade students were investigated to determine the degree to which these students changed answers, the extent to which they gained from changing, and the reasons they advanced for the changes. Students were from a predominantly suburban district (n=88) and a predominantly rural district (n=45). Relationships of answer changing behavior with test score, gender, and test-taking strategy were also examined. The majority of changes were wrong-to-right for all subgroups. High-scoring students changed fewer answers but had a higher proportion of changes from wrong-to-right than did other students. Reasons for changing, determined through interviews, were similar to those given by adults in previous studies, but with more clerical changes. For each test-taking strategy, students who changed tended to gain. Students should not be discouraged from changing answers on tests, especially given the preponderance of gain from changing, even for subgroups. Three tables present study data and there is a 21-item list of references. (Author/SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED342803

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

RUDOLPH V. RICHICHI

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Why Do Young Students Change Answers on Tests?

Robert F. McMorris, Shirley P. Schwarz, Rudoloph V. Richichi,

Michael Fischer, Nancy M. Buczek, C. Lynn Chevalier,

Kelli A. Meland

State University of New York at Albany

Running Head: WHY CHANGE

BEST COPY AVAILABLE

TM01 794

Abstract

Effects of answer changing on tests by fifth- and sixth-grade students were investigated to determine the degree to which these students changed answers, the extent to which they gained from changing, and the reasons they advanced for the changes. Relationships of answer changing behavior with test score, gender, and test-taking strategy were also examined. The majority of changes were wrong-to-right for all subgroups. High-scoring students changed fewer answers but had a higher proportion of changes from wrong-to-right than did other students. Reasons for changing were similar to those given by adults in previous studies but with more clerical changes. For each test-taking strategy, students who changed tended to gain. Students should not be discouraged from changing answers on tests, especially given the preponderance of gain from changing even for subgroups defined using various dimensions.

Why Do Young Students Change Answers on Tests?

Should students change answers on objective tests if they have a good reason for it? Or should they retain their initial answers because their first impression may have been correct, and answer changing would only lower their scores? Over six decades, more than three dozen published studies have been conducted, primarily with college students, trying to rebuke as myth the belief in the accuracy of first impression. The unanimous research finding has been that the majority of students gain from answer changing.

A brief literature review was organized to address the following questions:

- How much answer changing is likely to occur and does it result in gain?
- Does answer changing behavior relate to subject variables such as performance level or gender?
- For what reasons do people change answers?
- Do younger students also gain from answer changing?

Answer changing. Most students (84%) across studies were found to change answers on tests; however, the median percentage of items changed was small (3.3%), with a range of 2.2% to approximately 9% (Benjamin, Cavell, & Shallenberger, 1984; Mueller & Wasser, 1977; Schwarz, McMorris, & DeMers, 1991). Nevertheless, the majority of students did gain from such changes.

Approximately two out of three changes in a true/false test format and three out of four changes in a multiple choice format were found to be from the wrong to the right alternative (Beck, 1978; Benjamin et al., 1984; Best, 1979; Johnston, 1978; Lehman, 1928; Mueller & Shwedel, 1975; Smith, White, & Coop, 1979).

Student performance level. Percentage of answers changed was negatively related to test performance. High performers were less likely to revise their answers (Best, 1979; Johnston, 1978; Lynch & Smith, 1975; Mueller & Shwedel, 1975; Sitton, Adams, & Anderson, 1980; Stoffer, Davis, & Brown, 1977). But when they did make changes, they were more likely to gain from those changes (McMorris, DeMers, & Schwarz, 1987; Mueller & Shwedel, 1975). For instance, McMorris et al. (1987) obtained a correlation of $r = .81$ between percentage of gain and percentage of change for the highest third performers as opposed to $r = .20$ for the lowest third performers.

Gender. Some researchers have examined the relationship between gender and frequency of answer changing. Females were generally found to change slightly more answers; this result was statistically significant in only one study (Skinner, 1983). Moreover, evidence concerning the relationship between sex and gain from answer changing is conflicting and ambiguous (Benjamin et al., 1984; McMorris et al., 1987). For instance, one pair of researchers reported that females gain less from answer changing (Reile & Briggs, 1952), while another researcher reported that females gain more (Bath, 1967).

Reasons for changing answers. Reasons for answer changing (McMorris et al.,1987; McMorris & Weideman, 1986; Schwarz et al., 1991; Shatz & Best, 1987) included: (a) Rereading and better understanding the question; (b) Rethinking and conceptualizing a better answer; (c) Learning from a later item; (d) Remembering more information; (e) Clues in the item; (f) Clerical corrections; and (g) Guessing. It was consistently found that when answers were changed for any logical reason other than guessing, the changes were generally from the wrong to the right alternative (56 to 72%, depending on the study), whereas if guessing was the only reason for answer changing, changes were as likely to be (a) from a wrong to another wrong alternative, (b) from a right to a wrong alternative, or (c) from a wrong to a right alternative (Shatz & Best, 1987). Ramsey, Ramsey, & Barnes (1987) found the most common reasons for answer changing on a statistical test were "changing one's mind" (49%) and misreading or misinterpretation of question (32%), with only 8% of changes involving clerical errors. Gain averaged .56 for misread and misinterpretation of question, .27 for changing one's mind, and .18 for clerical errors. Finally, McMorris and Weideman (1986) found a preponderance of "thoughtful" as opposed to "clerical" changes, and, in a later study, McMorris et al. (1987) found that "remembering" as a reason for changing an answer produced the highest percentage of gain.

Answer changing among younger students. Do the above findings replicate with younger children? Do younger students benefit to the same extent from answer changing as do college students?

Beck (1978) examined the answer changing behavior of third graders on the Metropolitan Reading Test. He found that although few changes were made (approximately 3%), two-thirds of the changes on the word knowledge section and over half of the changes in the reading comprehension section were wrong to right. Comparing the twenty highest to the twenty lowest achievers in terms of total reading test scores, Beck found the high achievers to make more changes, although this difference was not statistically significant. Nevertheless, the high achievers had a significantly higher number of wrong-to-right changes than did the low achievers. Finally, no statistically significant sex differences were obtained on either the number or type of change.

Crocker and Benson (1980) found seventh graders to be more reluctant to change their answers than were college students, with 46 to 57% of young students as compared to 84% of college students changing at least one answer. Nevertheless, the median percentage of answers changed was slightly higher for seventh graders (4%) than for the college population (3.3%). Furthermore, younger students gained more from those changes (three-to-one wrong-to-right versus right-to-wrong), especially when items involved the option, "I do not know."

Muller (1987) found a positive correlation between fifth, ninth, and twelfth graders' performance and percentage gain, with

high performers making a relatively higher percentage of wrong-to-right changes than did low performers. Furthermore, scores on unchanged items were found to correlate negatively with the amount of answer changing of fifth, ninth, and twelfth graders. However, the above relationships reached statistical significance only for fifth graders.

The following questions remain:

- Can young students accurately evaluate when their original answer is incorrect? Are they aware of their reasons for answer changing? What are those reasons?
- Are they capable of discovering or remembering additional relevant information as they go through the test?
- Does answer-changing behavior relate to test score or to gender?
- Are each of several test-taking strategies conducive to gain?

Method

Fifth- and sixth-grade students from two school districts participated in the study. One district was predominantly suburban (n = 88), the other was predominantly rural (n = 45). Each school district gave written permission for the study, and students were given parental permission to participate in this research.

Instruments included two forms of a questionnaire and a structured interview. Form A of the questionnaire consisted of nine questions; questions 1-6 had three choices each on a Likert-type scale, and questions 7-9 were open ended. Form B of

the questionnaire duplicated the same questions except for questions 8 and 9 which provided a list of possible reasons for changing including an "other," fill-in-the-blank option. The response choices were taken from adults' responses in previous studies, such as "misread the question" or "a clerical error".

The structured answer-changing interview consisted of a maximum of twelve open-ended questions dealing especially with test-taking strategy and reasons for changing items on the test just taken. The interviewers were university graduate students doing advanced educational research; they were trained to administer the interview consistently. Each interviewer had a sheet of probes to prompt or help the student clarify his/her responses. The interviewer recorded the subject's responses. Interviewers' summaries were coded independently by two authors, with any discrepancies resolved either between the two coders or with the help of a third author.

The sixth-grade suburban students took a 33-item, teacher-made multiple-choice test of current events and subsequently were administered the questionnaire. Forty-two of the students were then interviewed by the researchers.

The fifth- and sixth-grade rural sample was comprised of a morning and an afternoon class. The morning class of sixth-grade students (n = 23) took a 25-item, teacher-made test of verbal ability. The afternoon class of fifth-grade students (n = 22) was examined on a 20-item science test based on its science textbook.

All subjects in the rural sample were subsequently administered a questionnaire and interviewed.

Results

On the questionnaire all students were asked about their answer changing beliefs. When asked whether they change answers on multiple choice tests, three-quarters of the students picked the middle option "sometimes." About two-thirds of the students thought their friends would get about the same mark by changing answers, with one-sixth thinking they would better their grade, and one-sixth thinking changing would make it worse. But when asked whether they personally would get a better mark by changing answers, 44% said "Yes," 40% said "about the same," and 16% chose "No, worse." Students apparently hadn't been misled about answer changing: 69% didn't think a teacher had advised against changing, and only 15% thought they had (16% weren't sure). Most (62%) thought they had had extra time to complete that day's test, 38% felt they had enough time, and no one indicated they had had insufficient time. Only 12% of the students would have changed more answers given more time.

Change and gain. Overall, 75% of the students changed at least one answer. The percentage of answers changed across items and people was 5.28.

Gain from changing was found for 56% of the students; 39% did not gain or lose, and 5% of the students lost. The number of changes from wrong-to-right (WR) exceeded the combined total of

wrong-to-wrong changes (WW) and right-to-wrong changes (RW). On the average, the percent gain from changing for an individual was 3.00.

Reasons for changing. The interviewers asked students why they changed answers on individual items. The replies were summarized in seven reasons as given in Table 1. The most popular reason was "rethought the answer"; virtually tied for second place were "guessed," "learned from a later item," and "clerical error."

The data were further analyzed according to students' test performance. Each class was divided into thirds based on their test scores prior to changing (i.e., test score minus WR plus RW); the high, middle, and low thirds from each of the classes were then combined for these analyses. The low third made more changes than did the middle or high thirds. "Rethought," "guessed," "clerical," and "misread" were reasons given more often by the low group than by either of the other groups. Only "remembered" was positively related to test performance.

Answer-changing behavior and test performance. Table 2 contains a summary of the effectiveness of answer changing for the high-, middle-, and low-scoring groups. For each group the total number of changes was delineated by type of change and described by summary measures of change and gain. Answer-changing behavior was further investigated using the correlation of percent gain (PGAIN) with percent change (PCHG), where PGAIN is defined as the number of RW changes subtracted from the number of WR changes

divided by the number of items on the test multiplied by 100, and PCHG is defined as the number of items changed divided by the number of items answered multiplied by 100.

The high group made the fewest number of item changes, the low group the most, but for all three groups the majority of changes were WR. The high group changed a larger proportion from wrong-to-right than did the middle or low groups, reflected in that group's greater gain per change (.70), although because they changed fewer items, their PGAIN mean was the lowest of the three group means. The correlation between PGAIN and PCHG was highest for the high group but was at least .5 for each of the groups. Also indicative of gain from changing, even for subgroups: the mean gain per change was close to .5 for the low group, and .55 overall.

Gender. A somewhat higher proportion of females changed answers than did males (82% vs. 69%). Similarly, the mean PCHG for females (5.56) slightly exceeded the mean PCHG for males (5.03).

Both males and females in this study tended to gain from answer changing, with males slightly more likely to gain (60% vs. 52%) and less likely to lose (1% vs. 10%). Similarly, the mean PGAIN for males (3.66) exceeded the mean PGAIN for females (2.15). Further, 32% of the males had at least a 6% gain, whereas only 18% of the females gained at least that much. Only one male and one female had a loss of 6%, the maximum loss for this

sample. Consistent with these findings, PGAIN was more highly related to PCHG for males (r = .82) than for females (r = .39).

Test-taking strategy. The test-taking strategy and corresponding PCHG AND PGAIN are shown in Table 3 for interviewed subjects from all groups. Excluding "Process of elimination"--mentioned as a strategy by three subjects, none of whom changed any items--all strategies produced change and gain in this sample. Again excluding "Process of elimination," the range for PCHG by strategy was narrow (6.7 - 4.0 = 2.7) as was the range for the average PGAIN (4.2 - 2.7 = 1.5).

Discussion

The basic consistencies of the prior research seem to be supported and extended by the results of this study. The students in this study tended to gain as a result of changing. The majority of changes were "wrong-to-right" for all groups. Further, the proportions of WR, WW, and RW changes were found to be similar to the proportions found for adults. The correlation of PGAIN with PCHG for all groups was .61, and the mean gain per change for all groups was .55, also supporting the conclusion that students who changed answers tended to gain. These young students seemed more optimistic about the outcomes of answer changing than have many previous samples (see Benjamin et al, 1984), and they had not apparently been taught the myth reported by many older students that they should "stick with their first answer."

Three-quarters of the students in this sample changed answers. The proportion of test takers who change answers, however, is partially a function of the number of items on which they are tested, so comparisons across studies of the proportions of test takers who change, and gain, are confounded. Statistics such as PCHG and PGAIN would likely be less affected by test length because of being per-item indices. To illustrate, in this study, 25% of the students did not change any answers, and an additional 14% did not change their total score even though they changed at least one item response. But the tests ranged from 20 to 33 items. In contrast, all students in McMorris and Weideman's (1986) study changed answers, with only 10% not changing their score by changing answers, but there were 121 total items. The difference in answer-changing behavior between the two studies seems less at the per-item level: PCHG and PGAIN were 5.28 and 3.00 in the current study, 6.90 and 3.00 in the 1986 study.

The subgroup in this study with the highest scores changed the fewest answers, perhaps partially due to their mastery of the subject material, which would be supported not only by their higher scores but also by their infrequent wrong-to-wrong changes. Such competence and relative freedom from WW changes are also consistent with the higher zero-order correlations between PGAIN and PCHG for the high group when compared to the middle or lower thirds. Further, the students with the highest scores tended to gain more per change, although even the lowest-scoring

students tended to gain. These findings are consistent with the trends in previous studies regarding the relationships of answer-changing behavior with test performance.

The present study may be seen as a response to the challenge by Benjamin et al. (1984, pp. 139,140) to determine why test takers change answers, and to seek the reasons for changing answers from a relatively understudied group, the younger students. This study provided evidence for the similarity of reasons for changing among children and adults, with perhaps more clerical errors noted here than in earlier research with older students. Still, the ratio of the number of changes for more thoughtful/cognitive reasons (rethought, learned, misread, remembered) to the number of guessing and clerical changes was almost 2:1 for these students.

When interviewed, the students seemed to use the same types of test-taking strategies as did adults (Schwarz et al., 1991). Further, for each strategy, students who changed answers tended to gain.

Several questions on answer-changing reasons and on test-taking strategies remain to be researched. For example, do individuals repeat the same reasons for changing across tests? Do reasons relate to specific items? (For example, are there specific items which many test takers change because they later remember the best answer?) Is test-taking strategy for individuals consistent across tests, either within or across

content areas? Although researchable questions remain, so does an implication for practice: Judging from the research results to date, test takers should not be discouraged from changing answers on tests, especially given the preponderance of gain from changing even for subgroups.

References

Bath, J. A. (1967). Answer-changing behavior on objective examinations. Journal of Educational Research, 61, 105-107.

Beck, M. D. (1978). The effect of item response changes on scores on an elementary reading achievement test. Journal of Educational Research, 71, 153-156.

Benjamin, L. T., Jr., Cavell, T. A., & Shallenberger, W. R., III. (1984). Staying with initial answers on objective tests: Is it a myth? Teaching of Psychology, 11, 133-141.

Best, J. B. (1979). Item difficulty and answer changing. Teaching of Psychology, 6, 228-230.

Crocker, L., & Benson, J. (1980). Does answer-changing affect test quality? Measurement and Evaluation in Guidance, 12, 233-239.

Johnston, J. J. (1978). Answer-changing behavior and grades. Teaching of Psychology, 5, 44-45.

Lehman, H. C. (1928). Does it pay to change initial decisions in a true-false test? School and Society, 28, 456-458.

Lynch, D. O., & Smith, B. C. (1975). Item response changes: Effects on test scores. Measurement and Evaluation in Guidance, 7, 220-224.

McMorris, R. F., DeMers, L. P., & Schwarz, S. P. (1987). Attitudes, behaviors, and reasons for changing responses following answer-changing instruction. Journal of Educational Measurement, 24, 131-143.

McMorris, R. F., & Weideman, A. H. (1986). Answer changing after instruction on answer changing. Measurement and Evaluation in Counseling and Development, 18, 93-101.

Mueller, D. J., & Shwedel, A. (1975). Some correlates of net gain resultant from answer changing on objective achievement test items. Journal of Educational Measurement, 12, 251-254.

Mueller, D. J., & Wasser, V. (1977). Implications of changing answers on objective test items. Journal of Educational Measurement, 14, 9-13.

Muller, E. W. (1987, April). Item response changes, grade and general ability. Paper presented at the meeting of the New England Research Organization, Stratton, VT.

Ramsey, P. H., Ramsey, P. P., & Barnes, M. J. (1987). Effects of student confidence and item difficulty on test score gains due to answer changing. Teaching of Psychology, 14, 206-210.

Reile, P. J., & Briggs, L. J. (1952). Should students change their initial answers on objective-type tests? More evidence regarding an old problem. Journal of Educational Psychology, 43, 110-115.

Schwarz, S. P., McMorris, R. F., & DeMers, L. P. (1991). Reasons for changing answers: An evaluation using personal interviews. Journal of Educational Measurement, 28, 163-171.

Shatz, M. A., & Best, J. B. (1987). Students' reasons for changing answers on objective tests. Teaching of Psychology, 14, 241-242.

Sitton, L. R., Adams, I. G., & Anderson, H. N. (1980). Personality correlates of students' patterns of changing answers on multiple-choice tests. Psychological Reports, 47, 655-660.

Skinner, N. F. (1983). Switching answers on multiple-choice questions: Shrewdness or shibboleth? Teaching of Psychology, 10, 220-222.

Smith, M., White, K. P., & Coop, R. H. (1979). The effect of item type on the consequences of changing answers on multiple-choice tests. Journal of Educational Measurement, 16, 203-208.

Stoffer, G. R., Davis, K. E., & Brown, J. B. (1977). The consequences of changing initial answers on objective tests: A stable effect and a stable misconception. Journal of Educational Research, 70, 272-277.

TABLE 1
Reasons for Changing Answers

Reason[a]	Test Performance[b] Low	Middle	High	Total
Rethought	17	10	10	37
Guessed	13	5	6	24
Learned	8	10	6	24
Clerical	14	3	5	22
Misread	8	4	2	14
Remembered	2	3	9	14
Other	2	5	0	7
Total	64	40	38	142

[a]The reasons were given to the interviewers by the 87 students interviewed.

[b]Each class was divided into thirds based on their scores prior to changing; the thirds were combined across classes.

TABLE 2
Answer-Changing Behavior

Answer Changing Index	Test Performance[a] Low f (%)	Middle f (%)	High f (%)	Total f (%)
Type of Change[b]				
WR	60 (62)	48 (70)	37 (80)	145 (68)
WW	24 (25)	10 (14)	4 (9)	38 (18)
RW	13 (13)	11 (16)	5 (11)	29 (14)
Total	97 (100)	69 (100)	46 (100)	212 (100)
Means[c]				
Chg/Indiv	2.2	1.6	1.0	1.6
PCHG	7.46	5.34	3.14	5.28
PGAIN	3.84	2.97	2.21	3.00
Gain/Chg	0.48	0.54	0.70	0.55
Correlation				
PCHG*PGAIN	0.512**	0.713**	0.796**	0.646**

[a]Each class was divided into thirds based on their scores prior to changing; here f = number of item changes.

[b]WR = wrong-to-right; WW = wrong-to-wrong; RW = right-to-wrong.

[c]Chg/Indiv = number of items changed per individual; PCHG = percent of items changed = number of items changed * 100 / number of items; PGAIN = percent gain = (WR - RW) * 100 / number of items; Gain/Chg = (WR - RW) / number of changes, i.e., expected gain per change.

** $p < .01$

TABLE 3
Test-Taking Strategy and Answer-Changing Behavior For the Interviewed Subjects

Strategy	n	PCHG[a]	PGAIN[a]
Leaves Blanks	19	4.6	2.7
Rereads None	19	4.9	3.0
Rereads Some	5	6.7	4.2
Rereads All	4	6.5	3.0
Leaves Blanks Rereads Some	23	4.2	3.4
Leaves Blanks Rereads All	12	4.1	2.8
Process of Elimination[b]	3	0.0	0.0
Other	2	4.0	4.0
Total	87	4.5	3.0

[a]PCHG = percent of items changed = number of items changed * 100 / number of items. PGAIN = percent gain = (WR - RW) * 100 / number of items.

[b]None of the three people specifying this strategy changed items.